JOURNAL OF APPLIED BEHAVIOR ANALYSIS 1996, 29, 93–98 NUMBER 1 (SPRING 1996)

# ON THE FUNCTION OF SELF-RESTRAINT AND ITS RELATIONSHIP TO SELF-INJURY

WAYNE W. FISHER
KENNEDY KRIEGER INSTITUTE AND JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
AND
BRIAN A. IWATA
UNIVERSITY OF FLORIDA

Three reports in this issue of *JABA* (Derby, Fisher, & Piazza, 1996; Fisher, Grace, & Murphy, 1996; Smith, Lerman, & Iwata, 1996) extend the results of a small group of studies on self-restraint, a curious phenomenon that has been observed as a correlate of self-injurious behavior (SIB). Self-restraint appears to be a behavior that is unique to individuals who engage in SIB. That is, we have not found any published reports of individuals who self-restrain but who do not engage in SIB, or of individuals who self-restrain in conjunction with other behavior problems (e.g., aggression). This raises some interesting questions about the relationship between self-restraint and SIB: Are both responses maintained by the same reinforcement contingency? Are the responses functionally independent but related through some historical process (e.g., correlated but different contingencies of reinforcement)? Does access to one response (or its termination) serve as reinforcement for the other? In reviewing some of the research on self-restraint, Smith, Iwata, Vollmer, and Pace (1992) found tentative support for each of these three hypotheses, and their data suggested that, as is the case with SIB, self-restraint may be influenced through several mechanisms of control. Each of the studies in the present series demonstrated a clear functional relationship between SIB and self-restraint by focusing on a single operant mechanism (e.g., assessing whether contingent access to self-restraint serves as reinforcement for SIB). In what follows, we will expand on these relationships, integrate findings from previous and current studies, and suggest directions for future research.

Address correspondence to Wayne W. Fisher, Neurobehavioral Unit, Kennedy Krieger Institute, 707 N. Broadway, Baltimore, Maryland 21205.

## *SIB and Self-Restraint as Members of the Same Functional Response Class*

One possible explanation of the correlation between SIB and self-restraint is that both are members of the same functional response class. Of the possible relationships between SIB and self-restraint outlined above, this account may have the most intuitive appeal, especially if both behaviors are maintained by social positive or negative reinforcement (e.g., attention or escape from demands). Iwata et al. (1994) have shown that SIB is maintained by social positive reinforcement in approximately 26% of cases and social negative reinforcement in about 38%, and suggested that SIB is maintained by automatic reinforcement (e.g., sensory stimulation) in another 26% of cases. The most common type of social positive reinforcement for SIB is attention in the form of a verbal reprimand (e.g., "Don't do that, you'll hurt yourself"). It is easy to envision how caregivers who provide verbal reprimands following SIB also might deliver attention in the form of praise contingent upon self-restraint (e.g., "Nice job not hit-

ting yourself"), a response that is incompatible with SIB. And if contingent attention maintains an individual's SIB, it might similarly function as reinforcement for self-restraint. The study by Derby et al. (1996) provides an example of SIB and self-restraint maintained by contingent attention.

The most common form of social negative reinforcement for SIB is avoidance of or escape from task demands. Given that it is often laborious to block or prevent self-restraint, it is not difficult to see how caregivers might present few or no demands to an individual engaged in this response. Thus, when escape from tasks is an effective reinforcer for SIB, it may similarly affect self-restraint. The data presented for 1 subject in the Smith et al. (1996) investigation indicated that SIB was maintained by escape from tasks and suggested that self-restraint was as well.

When SIB is maintained by automatic reinforcement, it is somewhat less clear how an individual might develop self-restraint maintained by the same or similar consequences. For example, if SIB were maintained by the sensory stimulation it produced, it is unlikely that self-restraint would result in the same or a similar form of stimulation.

### *SIB and Self-Restraint as Functionally Independent Responses*

Topographically distinct forms of aberrant behavior (e.g., aggression, SIB) may be maintained by the same consequences (i.e., forming a functional response class) or by different ones (i.e., functionally independent; Derby et al., 1994). But given that both SIB and self-restraint are rare, it is extremely unlikely that by chance alone the two responses would be acquired and maintained independently by distinct contingencies. It may be more plausible that SIB is acquired first and is maintained by one contingency (e.g., escape). Caregivers subsequently may provide a different consequence that maintains self-restraint (e.g., attention) in order to lessen the probability of SIB. Although there are no clear examples of such a relationship between SIB and self-restraint in the literature, we have observed that caregivers encourage (e.g., verbally or physically prompt) self-restraint and provide praise following the initiation of this response. Thus, this remains an untested hypothesis, but one worthy of further investigation.

### *Access to Self-Restraint as Positive Reinforcement of SIB*

Self-restraint, when it occurs, is generally a high-probability response, and it is well established that contingent access to a high-probability response can function as reinforcement for another response (e.g., Konarski, Johnson, Crowell, & Whitman, 1980; Premack, 1963). However, it may not be readily apparent how a contingent relationship between SIB and self-restraint might develop in the natural environment. That is, why would the environment require an individual to engage in self-injury in order to gain access to self-restraint? One possibility is that when self-restraint interferes with adaptive behavior, caregivers may interrupt, discourage, or even punish this response if it occurs in the absence of self-injury. However, if a burst of SIB occurs immediately after self-restraint is interrupted, caregivers may allow or even encourage the client to resume self-restraint in order to terminate SIB. Thus, if self-restraint is discouraged (i.e., response deprivation) when SIB is absent but is allowed following SIB, then it is possible that contingent access to self-restraint (a high-probability response) serves as reinforcement for SIB. The investigation by Smith et al. (1996) demonstrates that self-restraint can function as reinforcement for SIB and illustrates a method for assessing this hypothesis.

*Termination of SIB as Negative Reinforcement of Self-Restraint*

With this account of self-restraint, regardless of what variables maintain SIB (e.g., attention, escape, sensory feedback), SIB also presumably produces negative consequences (i.e., pain). Thus, the negative reinforcement hypothesis of self-restraint maintains that individuals escape or avoid these unpleasant consequences by engaging in self-restraint, a response that is incompatible with SIB (e.g., Baroff & Tate, 1968; Rojahn, Mulick, McCoy, & Schroeder, 1978). At first glance, this account of self-restraint may appear to be somewhat implausible. (Why go to all the trouble of hitting and then self-restraining, when one could just stop hitting in the first place?)

One possible but untested explanation is that if SIB produces both favorable (e.g., attention) and aversive consequences (e.g., pain), the effects on SIB may be similar to those associated with approach–avoidance paradigms (e.g., Miller, 1959), in which the variables that affect the probability of a response are in direct competition (i.e., conflict). When a single response is correlated with both favorable and unfavorable consequences, the relative probability of that response occurring (i.e., approach) or not occurring (i.e., avoidance) is dependent upon a number of factors, including relative schedule densities, quality of the favorable consequence, aversiveness of the negative consequence, prior exposure to deprivation or satiation conditions, and so forth. For example, other things being equal (e.g., consequence quality), the response should (a) occur when the schedule is denser for the favorable (e.g., variable-ratio [VR] 5) than for the unfavorable (e.g., VR 20) consequence and (b) not occur when the schedule is leaner for the favorable consequence than for the unfavorable consequence. Similarly, given equal schedules for the favorable and unfavorable consequences (e.g., both VR 10), the response should (a) occur when the quality of the favorable consequence is high enough to override the aversiveness of the unfavorable consequence and (b) not occur when the quality of the favorable event is not sufficient to override the aversiveness of the unfavorable consequence.

It is unlikely that the consequences that concurrently increase (e.g., attention) and decrease (e.g., pain) the probability of SIB are constant over time. For example, the potency of attention as a reinforcer for SIB may increase following periods of deprivation and decrease after satiation (Vollmer & Iwata, 1991). Conversely, the negative effects of SIB may be greater at certain times (e.g., blows to the head may be more painful after the development of a contusion; alternatively, the initial responses in an episode may be more painful than subsequent responses). These factors (i.e., concurrent favorable and unfavorable consequences for SIB that vary in potency over time) may help to explain why SIB occurs at certain times and not at others, and why SIB sometimes is replaced by a response that is incompatible with it (i.e., self-restraint). That is, self-restraint may be what Terrace (1974) has labeled an *antagonistic* response that occurs when the favorable consequences of SIB (i.e., motivation to respond) are slightly overridden by the unfavorable consequences (i.e., motivation to inhibit responding).

Terrace (1974) has shown that for a short time following a change from a reinforcement schedule to extinction, individuals emit a second response that is incompatible with the response placed on extinction. This second response, referred to as an antagonistic response, may be similar to self-restraint. When a previously reinforced response is placed on extinction during discrimination training, it often increases temporarily (i.e., an extinction burst) and then gradually decreases. That is, the response continues to be

affected by the previous reinforcement contingency until the individual learns that the response no longer produces reinforcement. Terrace argued that emitting a response that was previously reinforced (i.e., nonreinforced responding) is aversive (see Terrace, 1972, for a discussion). He has provided rather compelling evidence for this position by showing that stimuli associated with nonreinforced responding during extinction become conditioned aversive stimuli (Terrace, 1971).

During this period following the change from reinforcement to extinction in discrimination training, the target response may be affected by both the previous reinforcement schedule and the current aversive schedule (due to nonreinforced responding). Thus, during this time there may be motivation to both respond and not respond (i.e., an approach–avoidance situation), and it is during this time that the antagonistic response occurs. The antagonistic response appears during the transition period when the target response is partially under the control of the previous reinforcement schedule and partially under the control of the aversive aspects of nonreinforced responding. By way of example, overeating may be replaced by throwing away the contents of the refrigerator (a response that is incompatible with eating), and it may occur early or late in the process of food consumption when motivation is greater for dieting than for eating.

Similarly, self-restraint as a clinical problem may be more likely to develop in individuals for whom the schedule or quality of reinforcement for SIB is often superseded by its negative effects. At other times, SIB occurs because its favorable consequences override the unfavorable ones. A response that is antagonistic to SIB (i.e., self-restraint) may be most probable when factors that increase the probability of SIB (e.g., attention, escape) are present but are slightly overridden by factors that decrease the probability of SIB (e.g., pain).

A definitive test of the negative reinforcement hypothesis of self-restraint is difficult (if not impossible) because the presumed reinforcer (i.e., escape from or avoidance of SIB) is an automatic consequence of self-restraint. For self-restraint to be on extinction, SIB (the presumed aversive event or establishing operation) would have to occur in the presence of self-restraint, a condition that cannot be experimentally controlled. If avoidance of SIB is the reinforcer for self-restraint and the two responses are incompatible, then it is not possible to place self-restraint on extinction (i.e., when self-restraint occurs, SIB cannot occur because the responses are mutually exclusive). Nevertheless, several clinical studies have provided indirect empirical support for the negative reinforcement hypothesis. One approach has been to manipulate the extent to which SIB may produce painful consequences (e.g., Fisher et al., 1996; Silverman, Watanabe, Marshall, & Baer, 1984), and a second has been to manipulate the availability of restraint during a functional analysis of SIB (Smith et al., 1992).

In these investigations, authors suggested that self-restraint was maintained by avoidance of SIB (a) when subjects displayed relatively high levels of self-restraint and low levels of SIB in conditions in which both responses were freely available (Fisher et al., 1996; Silverman et al., 1984; Smith et al., 1992); (b) SIB increased immediately when subjects were exposed to conditions in which self-restraint was unavailable or restricted (Fisher et al.; Smith et al.); and (c) self-restraint decreased when procedures were implemented to block or attenuate the negative consequences of SIB (Fisher et al.; Silverman et al.).

Thus empirical data from both basic and applied investigations provide support for (but not proof of) the hypothesis that self-

restraint can be maintained by escape from the aversive consequences of SIB. Results from basic research suggest that a response that is incompatible with (or antagonistic to) a target response may appear when the target response is simultaneously affected by appetitive and aversive schedules (Miller, 1959; Terrace, 1974). Clinical research findings suggest that self-restraint may represent a real-world example of an antagonistic response, which develops into a clinical problem almost exclusively among individuals with SIB, because SIB can be simultaneously and chronically affected by appetitive and aversive schedules (e.g., contingent attention and pain).

### *Clinical Implications and Suggestions for Future Research*

Given the number of possible functions of and relationships between SIB and self-restraint, it may not be practical or possible to test all possible combinations (e.g., attention-maintained SIB might be accompanied by self-restraint maintained by contingent attention or tangible items, escape from demands or SIB, or automatic reinforcement). Thus, a reasonable approach to clinical assessment and research on SIB and self-restraint might be first to use the procedures described by Smith et al. (1992) (a) to determine the function of SIB and (b) to generate hypotheses about the function of self-restraint. With this method, self-restraint is alternately allowed or prevented during a functional analysis of SIB. These hypotheses could then form the basis of specific subsequent analyses on the function of self-restraint and its relationship to SIB, as was done in the three accompanying reports (Derby et al., 1996; Fisher et al., 1996; Smith et al., 1996).

When SIB is determined to be maintained by a specific form of social reinforcement (e.g., attention), it may be reasonable to hypothesize that self-restraint is maintained by the same consequence, especially if higher levels of self-restraint occur in the condition associated with increased SIB (Smith et al., 1992). Then, a specific test of this hypothesis could be conducted using methods similar to those described by Derby et al. (1996). If it is observed during the functional analysis that SIB is occasioned by interruption or prevention of self-restraint, then it may be reasonable to hypothesize that SIB is maintained by access to self-restraint and to test this hypothesis using the methods described by Smith et al. (1996). Assessments designed to indirectly evaluate the hypothesis that self-restraint is maintained by avoidance of SIB (e.g., Fisher et al., 1996; Silverman et al., 1984) may be most appropriate for individuals who display high levels of self-restraint and low levels of SIB across functional analysis conditions when self-restraint is available relative to when it is unavailable. Finally, additional basic and applied research is needed on the development of competing or antagonistic responses when target responses are concurrently associated with appetitive and aversive consequences (i.e., approach–avoidance paradigms) to test the generality of the phenomenon observed by Terrace (1974).

## REFERENCES

Baroff, G. S., & Tate, B. G. (1968). The use of aversive stimulation in the treatment of chronic self-injurious behavior. *Journal of the American Academy of Child Psychiatry, 7,* 454–470.

Derby, K. M., Fisher, W. W., & Piazza, C. C. (1996). The effects of contingent and noncontingent attention on self-injury and self-restraint. *Journal of Applied Behavior Analysis, 29,* 107–110.

Derby, K. M., Wacker, D. P., Peck, S., Sasso, G., DeRaad, A., Berg, W., Asmus, J., & Sonya, S. (1994). Functional analysis of separate topographies of aberrant behavior. *Journal of Applied Behavior Analysis, 27,* 267–278.

Fisher, W. W., Grace, N. C., & Murphy, C. (1996). Further analysis of the relationship between self-injury and self-restraint. *Journal of Applied Behavior Analysis, 29,* 103–106.

Iwata, B. A., Pace, G. M., Dorsey, M. F., Zarcone, J.

R., Vollmer, T. R., Smith, R. G., Rodgers, T. A., Lerman, D. C., Shore, B. A., Mazaleski, J. L., Goh, H. L., Cowdery, G. E., Kalsher, M. J., McCosh, K. C., & Willis, K. D. (1994). The functions of self-injurious behavior: An experimental-epidemiological analysis. *Journal of Applied Behavior Analysis, 27,* 215–240.

Konarski, E. A., Johnson, M. R., Crowell, C. R., & Whitman, T. I. (1980). Response deprivation and reinforcement in applied settings: A preliminary analysis. *Journal of Applied Behavior Analysis, 13,* 595–609.

Miller, N. E. (1959). Liberalization of basic S-R concepts: Extensions to conflict behavior, motivation, and social learning. In S. Koch (Ed.), *Psychology: A study of a science* (Vol. 2, pp. 196–293). New York: McGraw-Hill.

Premack, D. (1963). Rate differential reinforcement in monkey manipulation. *Journal of the Experimental Analysis of Behavior, 6,* 81–89.

Rojahn, J., Mulick, J. A., McCoy, D., & Schroeder, S. R. (1978). Setting events, adaptive clothing, and the modification of head banging and self-restraint in two profoundly retarded adults. *Behavior Analysis and Modification, 2,* 185–196.

Silverman, K., Watanabe, K., Marshall, A. M., & Baer, D. M. (1984). Reducing self-injury and corresponding self-restraint through the strategic use of protective clothing. *Journal of Applied Behavior Analysis, 17,* 545–552.

Smith, R. G., Iwata, B. A., Vollmer, T. R., & Pace, G. M. (1992). On the relationship between self-injurious behavior and self-restraint. *Journal of Applied Behavior Analysis, 25,* 433–445.

Smith, R. G., Lerman, D. C., & Iwata, B. A. (1996). Self-restraint as positive reinforcement for self-injurious behavior. *Journal of Applied Behavior Analysis, 29,* 99–102.

Terrace, H. S. (1971). Escape from S−. *Learning and Motivation, 2,* 148–163.

Terrace, H. S. (1972). Inhibition and discrimination. In R. A. Boakes & M. S. Halliday (Eds.), *Inhibition and learning* (pp. 301–336). New York: Academic Press.

Terrace, H. S. (1974). On the nature of nonresponding in discrimination learning with and without errors. *Journal of the Experimental Analysis of Behavior, 22,* 151–159.

Vollmer, T. R., & Iwata, B. A. (1991). Establishing operations and reinforcement effects. *Journal of Applied Behavior Analysis, 24,* 279–291.